



Universidade de Brasília
Instituto de Química

Comparação da performance de três espectrômetros portáteis de infravermelho próximo para identificação de madeira mogno e espécies similares

PRISCILA VERAS DOS ANJOS LOPES

TRABALHO DE CONCLUSÃO DE CURSO

Orientadora: Dra. Fernanda Piccolo Pieruzzi
Coorientador: Dr. Jez William Batista Braga

Brasília, 2019

Comparação da performance de três espectrômetros portáteis iguais de infravermelho próximo visando utilização em campo para identificação de madeira mogno e espécies similares

Trabalho de conclusão de curso apresentado para a obtenção do título de bacharel em de Química Tecnológica da Universidade de Brasília.

Orientadora: Dra. Fernanda Piccolo Pieruzzi

Brasília, 2019

AGRADECIMENTOS

A Deus, o Senhor da minha vida, pela oportunidade de realizar esse grande passo profissional, sempre me conduzindo no caminho do respeito e do amor ao próximo. Graças a Ele, tenho muitas pessoas que me deram força e não me deixaram desistir.

Agradeço a minha família amada, única e especial. Minha mãe Angélica, meu pai do coração Petrônio, minha irmã Thaís, minha avó Noêmia, minha sobrinha Maria, meu cunhado Luiz Antônio, meu irmão do coração Alexandre e meu pai Mario. Eles me deram apoio, alegrias, oportunidades e me ajudaram a ser quem eu sou hoje. Agradeço, também, ao João Paulo, por me acompanhar nessa longa caminhada e a minha família do coração meus tios Eduardo, Márcia, Carmem e Adelmá por estarem presente em todos os momentos importantes da minha vida. Não posso deixar de agradecer a todos os meus familiares que estão fisicamente a quilômetros daqui, mas sempre enviaram pensamentos positivos e torceram pelo meu sucesso. Ao: meu avô Pedro (*in memoriam*) e meu tios: Pierre, Péricles, Fábio, Rose, Cris e Douglas. As minhas primas e aos meus primos.

Foram muitas batalhas que enfrentei todos esses anos, e não poderia deixar de agradecer a todos os meus amigos por estarem comigo. Alguns desde a minha infância: Bruna, Alicia, Filipe, Thaís, Débora e Ana Caroline. Outros, chegaram há pouco tempo, mas já ocupam posição de destaque no meu coração Arthur, Danielli, Renata, Aninha, Ana Carmo, Andressa, Emanuelle, Luana e Marly.

Sou grata também a tantos olhares amigos que vejo na UnB, aos excelentes professores e equipe do Instituto de Química.

Agradeço ao meu melhor amigo Filipe por me ajudar a conseguir o estágio no Serviço Florestal Brasileiro e à Dra. Tereza Pastore, peça fundamental para chegar até esse trabalho. Ela me apresentou a minha orientadora: a maravilhosa, paciente e amiga Dra. Fernanda Pieruzzi, e ao maravilhoso e sereno mestre Dr. Jez Braga. Também tive a oportunidade de conhecer as pessoas maravilhosas do Laboratório de Produtos Florestais: Lúcia, Dhébora e em especial ao Alessandro pelos ensinamentos sobre quimiometria.

LISTA DE FIGURAS

Figura 1 – Amostras de madeira cortadas e lixadas, sendo a) e b) mogno (M), c) e d) cedrinho (CE), e) e f) curupixá (CX), g) e h) cedro (CD).....	19
Figura 2 - Espectrômetro portátil microPhazir™ RX Analyzer (Thermo Scientific).....	20
Figura 3 - Procedimento realizado na montagem dos modelos de discriminação.....	21
Figura 4 - Gráficos bidimensionais dos espectros de mogno a partir da PCA dos três equipamentos estudados, antes e após o pré-processamento dos espectros: (a) sem pré-processamento; (b) pré-processamento mean center; (c) pré-processamento primeira derivada com Savitzky-Golay e (d) junção dos pré-processamentos mean center e 1ª derivada.....	23
Figura 5 – Espectros brutos obtidos dos três equipamentos para as espécies: (a) mogno; (b) curupixá; (c) cedro e (d) cedrinho.....	25
Figura 6 – Espectros brutos obtidos dos três equipamentos após a aplicação da 1ª derivada com alisamento Savitzky-Golay das espécies: (a) mogno; (b) curupixá; (c) cedro e (d) cedrinho	26
Figura 7 – Distribuição dos valores estimados do conjunto de treinamento e validação para os modelos PLSR para o mogno. a) calibração eq. 1 com validação eq. 1; b) calibração eq. 3 com validação eq. 2; c) calibração modelo global com validação eq. 2; Classe 1 = mogno; Classe 1 = mogno; Classe 0 = demais espécies estudadas; LD = limite de discriminação; Val_outras = espectros de validação das demais espécies e Val_mogno = espectros de validação do mogno.....	31
Figura 8 – Espectros brutos de mogno obtidos com dois equipamentos em que: (a) aplicação da 1ª derivada com alisamento Savitzky Golay e (b) correção com EPO.....	33
Figura 9 - Gráficos bidimensionais dos espectros de mogno a partir da PCA de dois três equipamentos com o pré-processamento mean center e sem a aplicação da EPO: (a) gráfico com duas PCs e (b) gráfico com três PCs.....	33
Figura 10 - Gráficos bidimensionais dos espectros de mogno a partir da PCA de dois três equipamentos com o pré-processamento mean center e com a aplicação da EPO: (a) gráfico com duas PCs e (b) gráfico com três PCs	34
Figura 11 - Gráficos bidimensionais dos espectros de cedrinho.....	39
Figura 12 - Gráficos bidimensionais dos espectros de cedro.....	40
Figura 13 - Gráficos bidimensionais dos espectros de curupixá.....	41

LISTA DE TABELAS

Tabela 1 – Valores dos resultados das figuras de mérito dos modelos de discriminação.....	28
---	----

LISTA DE SÍMBOLOS, NOMENCLATURAS E ABREVIATÓES

CITES – Convenção sobre Comércio Internacional das Espécies da Flora e Fauna Selvagens em Perigo de Extinção (*Convention on International Trade in Endangered Species of Wild Fauna and Flora*)

EPO – Ortogonalização de parâmetros externos (*External Parameter Orthogonalization*)

LD – Limite de discriminação

LDA – Análise discriminante linear (*Linear Discriminant Analysis*)

LPF – Laboratório de Produtos Florestais

NIRS – Espectroscopia no infravermelho próximo (*Near Infrared Spectroscopy*)

PCA – Análise dos Componentes Principais (*Principal Component Analysis*)

PC – Componente Principal (*Principal Component*)

PLSR – Regressão por mínimos quadrados parciais (*Partial Least Squares Regression*)

PLS – Mínimos Quadrados Parciais (*Partial Least Squares*)

RMSEC – Erro médio quadrático de calibração (*Root Mean Square Error of Calibration*)

RMSECV – Erro médio quadrático da validação cruzada (*Root Mean Square Error of Cross-validation*)

SFB – Serviço Florestal Brasileiro

SNV – Padronização Normal de Sinval (*Standard Normal Variate*)

TEF – Taxa de eficiência

TFP – Taxa de falso positivo

TFN – Taxa de falso negativo

RESUMO

O mogno brasileiro, *Swietenia macrophylla* King, é uma espécie arbórea nativa da Amazônia ameaçada de extinção. Apresenta exploração restrita, de acordo com a legislação ambiental brasileira vigente. Devido a semelhança visual de sua madeira com a de outras espécies, o mogno por vezes é negociado como se fosse de outras espécies de legislação menos restritiva. A tecnologia NIR (*near infrared*), vem se mostrando como uma ferramenta promissora na análise da discriminação de espécies de madeira. Entretanto, para que a técnica possa ser utilizada para fins de fiscalização, faz-se necessário o carregamento dos modelos de discriminação em múltiplos equipamentos que, em campo, devem funcionar com a mesma eficiência. Assim, o objetivo principal deste trabalho é a comparação das respostas espectrais e o desempenho na discriminação da madeira de *S. macrophylla* de espécies visualmente semelhantes a partir de espectros de três espectrômetros NIR portáteis de mesma marca e modelo. Para o estudo, foram utilizadas madeiras de *Cedrela odorata* L. (cedro), *Micropolis melinoniana* (curupixá) e *Erisma uncinatum* (cedrinho), além da *S. macrophylla*. Os espectros obtidos em cada equipamento foram comparados entre si pela análise por componentes principais (PCA). A discriminação das espécies foi realizada a partir de modelos de discriminação por mínimo quadrados parciais (PLSR). O treinamento de cada modelo teve sua eficiência testada a partir da validação tanto de espectros obtidos com o mesmo equipamento quanto por espectros dos demais equipamentos. A partir da PCA, observou-se a distribuição dos espectros em agrupamentos distintos para cada equipamento, com maior separação entre o grupo de espectros do equipamento 1 em relação aos equipamentos 2 e 3. Em relação aos modelos de discriminação das espécies, aqueles desenvolvidos a partir de amostras de calibração e validação originadas de um mesmo equipamento tenderam a obter taxas de eficiência (TEF) acima de 95%. Ao se validar o modelo de calibração com espectros originados de outro equipamento, as TEFs tenderam a uma queda drástica, entre 0 e 67,5%. A construção de modelos utilizando espectros obtidos nos três equipamentos durante a fase de treinamento, permitiu com que a validação com espectros de diferentes origens apresentasse TEFs entre 70 a 100%. A aplicação de uma correção preliminar baseada na ortogonalização de parâmetros externos (EPO) permitiu uma aproximação das respostas espectrais dos equipamentos analisados. Conclui-se que existem diferenças espectrais entre os equipamentos, podendo ser minimizadas pela utilização de ferramentas como a EPO, cuja aplicação é uma perspectiva de continuidade deste trabalho.

Palavras-chave: mogno, tecnologia NIR, espectrômetro NIR portátil, PCA, PLSR.

ABSTRACT

The Brazilian mahogany, *Swietenia macrophylla* King, is an endangered native species from Amazon under restricted exploitation according to the current Brazilian legislation. Due to the wood similarity with other species, its timber is sometimes traded as another species under less restrictive legislation. The NIR technology (Near Infrared) has been proving to be a promising tool in the analysis of wood species discrimination. However, the utilization of this technique during inspection activities requires that the discrimination models loaded on multiple equipment works properly, with same efficiency in the field. Thus, the main goal of this work is the comparison of the spectral responses and the discriminating models of *S. Macrophylla* wood from visually similar species from spectra obtained in three portable NIR equipment of same brand and model. For this study, samples of each wood specie as *Cedrela odorata* L. (Cedar), *Micropolis melinoniana* (Curupixá), *Erisma uncinatum* (Cedrinho), and *S. macrophylla* were used. The results obtained in each equipment were compared to each other by Principal Component Analysis (PCA). The species discrimination was performed from Partial Least Squares Regression (PLSR). The training of each model was tested by the validation with spectra obtained from the same equipment and the spectra obtained with other equipment. The PCA highlighted a spectrum distribution in distinct groups. The variation was higher between the equipment 1 comparing 2 and 3. Regarding the discrimination models of the species, those that were developed from calibration and validation samples of the same equipment tended to obtain efficiency rates (ER) above 95%. The ER tended to a drastic drop between 0 and 67,5% when the calibration model was validated with spectra obtained with other equipment, as expected according to the spectral differences observed between the equipment. The models whose trainings were made by mixing spectra obtained in the three equipments allowed ER between 70 to 100% after validation with spectra of different origins. The External Parameter Orthogonalization (EPO) application allowed a closer spectra response between the equipments analyzed. It can be concluded that there are spectral differences between the equipments even though they belong to the same brand and model or if they have the same time of use. The preliminary analyzes shows EPO as a promising tool on the attenuation of the differences between equipments and its application is a perspective of continuity of this work.

Keywords: mahogany, NIR technology, NIR spectrometer, PCA, PLSR

SUMÁRIO

1 INTRODUÇÃO.....	10
2 OBJETIVO.....	12
2.1 Objetivo Geral.....	12
3 REFERENCIAL TEÓRICO	13
3.1 Mogno (<i>Swietenia Macrophylla King</i>).....	13
3.2 Tecnologia NIR (<i>Near Infrared Technology</i>).....	14
3.3 Quimiometria.....	15
3.3.1 Análise dos componentes principais (PCA).....	15
3.3.2 Pré-processamentos e ortogonalização de parâmetros externos (EPO).....	16
3.3.3 Regressão por mínimos quadrados parciais (PLSR).....	17
4 MATERIAIS E MÉTODOS	19
4.1 Espécies de madeiras estudadas	19
4.2 Preparação do material.....	19
4.3 Aquisição dos espectros NIR.....	20
4.4 Processamento de espectros e construção de modelos quimimétricos.....	20
5 RESULTADOS E DISCUSSÕES	23
5.1 Análise exploratória dos espectros obtidos com os três equipamentos empregando PCA.....	23
5.2 Modelos de predição.....	27
5.3 Avaliação preliminar da correção da diferença entre os equipamentos por EPO.....	32
6 CONCLUSÃO E PERSPECTIVAS	35
7 REFERÊNCIAS BIBLIOGRÁFICAS	36
8 ANEXO.....	39

1 INTRODUÇÃO

O mogno brasileiro, *Swietenia macrophylla* King, é uma espécie arbórea típica da Amazônia, mas que também ocorre em estados do nordeste e centro-oeste do país. A espécie apresenta madeira de alta qualidade com ampla gama de aplicações, desde construção civil, móveis e até instrumentos musicais (SOUZA; CAMARGOS, 2014). Essa espécie apresenta restrição na exploração e controle do seu comércio por ter sofrido uma exploração extensiva ao longo dos anos e por estar na lista de espécies da flora brasileira ameaçada de extinção (BRASIL, 2014).

Por conta da semelhança visual com outras espécies e na tentativa de se burlar a legislação ambiental brasileira vigente, a madeira de *S. macrophylla*, por vezes, é negociada como se fosse de outras espécies, ou seja, sob nomes diferentes. Por outro lado, algumas madeiras semelhantes pertencentes à família Meliaceae são comercializadas como "mogno" ou como sendo de "padrão mogno". Pode-se citar as espécies semelhantes ao mogno, como *Cedrela odorata* L. (cedro) e *Erismia uncinatum* (cedrinho) (PASTORE et al., 2011).

Tradicionalmente, a identificação da madeira é feita por anatomistas especializados, que comparam as amostras a serem identificadas com aquelas depositadas em bibliotecas de referência a partir da observação (em lupa ou microscópio) dos caracteres morfoanatômicos da madeira. Por requerer um longo tempo de treinamento, ferramentas auxiliares à tradicional identificação anatômica e que possam fornecer resultados seguros vem sendo desenvolvidos pelo Laboratório de Produtos Florestais (LPF).

Dentre elas, a técnica de identificação da madeira a partir de estudos de espectroscopia no infravermelho próximo (NIR) associada à Quimiometria vem se mostrando uma ferramenta promissora, sendo capaz de discriminar amostras de madeira não apenas entre espécies, mas também entre diferentes locais de origem para uma mesma espécie (BERGO et al., 2016; DE MUNIZ et al., 2012; OLIVEIRA et al., 2015; PASTORE et al., 2011; SILVA et al., 2018; SOARES et al., 2017).

A técnica NIR se baseia na aquisição direta de medidas de reflectância, que carregam informação dos diversos grupos funcionais presentes nas moléculas de alto peso da madeira, tais como celulose, hemicelulose, lignina e de menor peso molecular (extrativos) (TSUCHIKAWA, 2007). O espectro de NIR da madeira é constituído de um conjunto de bandas formadas pela sobreposição de várias transições vibracionais na região

de sobretons e combinações de bandas que trazem informações dos constituintes químicos da madeira como se fossem uma impressão digital da espécie (PASQUINI, 2003).

Esta técnica destaca-se por ser rápida, não destrutiva, reprodutível, precisa e requer o mínimo preparo da amostra mas, necessita da formação de um modelo matemático que relacione os espectros obtidos, com uma ou mais propriedades de interesse, de maneira quantitativa ou qualitativa (PASQUINI, 2003).

Assim, surgiu a necessidade da utilização de ferramentas estatísticas e matemáticas das quais podem se destacar: a análise de componentes principais (*PCA*, do inglês *Principal Component Analysis*), regressão por mínimos quadrados parciais (*PLSR*, do inglês *Partial Least Squares Regression*) e análise discriminante linear (*LDA*, do inglês *Linear Discriminant Analysis*) (SO et al., 2004).

Entretanto, tais modelos matemáticos devem ser construídos de forma que possam ser carregados em múltiplos equipamentos que, por sua vez, em campo, devem funcionar com a mesma eficiência.

2 OBJETIVO

O objetivo principal deste trabalho foi a avaliação da resposta espectral e do desempenho na discriminação da madeira de *S. macrophylla* de espécies visualmente semelhantes a partir de três equipamentos NIR portáteis de mesma marca e modelo.

2.1 Objetivos Específicos

- I. Comparar os espectros de madeira de *S. macrophylla* (mogno); *C. odorata* (cedro), *M. melinoniana* (curupixá) e *E. uncinatum* (cedrinho) obtidos a partir de três equipamentos MicroPhazir;
- II. Desenvolver modelos de identificação/discriminação das madeiras de *S. macrophylla* e similares a partir de espectros de três diferentes equipamentos MicroPhazir;
- III. Verificar se há perdas na eficiência nos modelos de identificação/discriminação ao se validar a calibração de um equipamento com espectros obtidos em outro equipamento de mesma marca e modelo;
- IV. Propor soluções que possam minimizar possíveis diferenças espectrais entre os equipamentos MicroPhazir, permitindo a transferência de calibração entre eles e evitando assim, a perda de eficiência de modelo de discriminação quando carregado em equipamento diferente daquele a partir do qual foram obtidos os espectros de calibração.

3 REFERENCIAL TEÓRICO

3.1 Mogno (*Swietenia macrophylla king*)

A espécie *Swietenia Macrophylla* King, pertencente à família da *Meliaceae*, é considerada como o “verdadeiro mogno”. Sua árvore pode medir aproximadamente 30 metros de altura e até mais de um metro de diâmetro. A madeira de mogno possui uma cor rosada-avermelhada, densidade média de 0,50 g/cm³, resistência contra fungos e insetos, é fácil de se trabalhar com ferramentas manuais ou elétricas e é considerada uma das madeiras mais caras do mundo para móveis finos (CORADIN; CAMARGOS, J. A. A. MARQUES; SILVA JÚNIOR, 2009; SOUZA; CAMARGOS, 2014)

A madeira do mogno foi intensamente explorada até 1994, época em que sua exploração foi restringida a apenas sob a forma de manejo florestal sustentável (Brasil, 1994). A convenção sobre o comércio internacional das espécies da flora e fauna selvagens em perigo de extinção (CITES) foi assinada pelo Brasil em 1975 para regular de forma eficaz o comércio de espécies da fauna e flora, prevenindo-as do perigo de extinção (IBAMA, 2018). Em 1995, o mogno passou a ter seu comércio internacional monitorado a partir de sua inclusão no anexo I da CITES e em 2002 foi incluída em sua lista do Anexo II no qual se incluem as espécies em perigo de extinção.

De acordo com a legislação brasileira, pelo decreto Nº 4.722 de 5 de junho de 2003, a exploração da espécie mogno em florestas nativas, primitivas ou regeneradas, somente será permitida sob a forma de manejo florestal sustentável, observando o prazo previsto no Decreto Nº 4.593, de 13 de fevereiro de 2003 (BRASIL, 2003). O decreto Nº 6.675, de 5 de junho de 2008, alterou o art. 3º do Decreto Nº 4.722 e passou a vigorar a proibição do abate de árvores da espécie *Swietenia macrophylla* (mogno), inclusive em áreas nas quais seja autorizada a supressão de vegetação (BRASIL, 2008).

Como consequência dessa situação, o mogno tem sido substituído por madeiras com características e aparência similares que geralmente são confundidas com mogno. Diante disso, houve a necessidade para que órgãos de fiscalização e técnicos que atuam no manejo florestal tenham acesso a informações que auxiliem no reconhecimento dessa espécie e que possam diferenciá-las de outras. (MACIEIRA et al., 2013; PASTORE et al., 2011).

3.2 Tecnologia NIR (*Near Infrared Technology*)

A região espectral do infravermelho próximo estende-se desde o limite superior da região de comprimento de onda da região visível (780 nm) até 2.500 nm. A espectroscopia NIR pode ser compreendida como uma modalidade de espectroscopia vibracional de energia relativamente alta. A região NIR do espectro eletromagnético permite observar os fenômenos de transferência de energia eletromagnética da radiação NIR para os sistemas vibracionais mecânicos atômicos encontrados principalmente em moléculas orgânicas que apresentam os grupos funcionais C-H, N-H, O-H e S-H (PASQUINI, 2003) .

A tecnologia NIR passou a ser rapidamente desenvolvida desde meados da década de 1970, tendo como suporte os fundamentos da espectroscopia vibracional, instrumentos espectrofotométricos NIR e a quimiometria (TIBOLA et al., 2018). Os usos mais importantes da radiação no infravermelho próximo são para a determinação quantitativa de espécies como água, proteínas, hidrocarbonetos de baixo peso molecular e gorduras. Essa ferramenta se tornou de grande importância em determinações quantitativas e o uso dessa técnica tem se expandido para determinar proteínas, óleos, umidade, lipídeos e celulose. As medidas podem ser realizadas através de sinais de reflexão, transmissão e reflectância difusa. No método da reflectância difusa, a radiação penetra na camada superficial das partículas, excitando modos vibracionais da molécula da amostra, que é, então, espalhada em todas as direções. A grande vantagem desse método é a rapidez e a simplicidade na preparação da amostra (SKOOG; HOLLER, F. JAMES NIEMAN, [s.d.]).

Atualmente, a espectroscopia NIR tem sido frequentemente aplicada como método analítico que fornece resultados eficientes para a determinação de moléculas orgânicas e variáveis qualitativas. Esta técnica está presente em praticamente todas as áreas devido ao nível de desenvolvimento que a tecnologia atingiu nos últimos anos (DE MUNIZ et al., 2012).

A combinação da tecnologia NIR com métodos quimiométricos aumentou a aplicabilidade para estudos da madeira, possibilitando a obtenção de mais informações a partir de seus espectros (BERGO, 2014).

3.3 Quimiometria

Com o avanço tecnológico, a chegada dos computadores e a sofisticação crescente das técnicas instrumentais, os métodos e análise de dados tiveram que se tornar mais sofisticados matematicamente e estatisticamente. A análise multivariada foi introduzida no tratamento de dados químicos e passou a se chamar de quimiometria. Em 1970, a

quimiometria foi formalizada como disciplina e linha de pesquisa e em 1974, o químico sueco Svante Wold e o matemático e químico estadunidense Bruce Kowalski fundaram a Sociedade Internacional da Quimiometria (ESBENSEN; GELADI, 1990; MARCELO M. DE SENA, RONEI J. POPPI, ROSA T.S. FRIGHETO, 1999).

A quimiometria é, então, a aplicação de métodos matemáticos e estatísticos usados para investigar, interpretar, classificar e prever conjuntos de dados de interesse químico. Esses métodos são utilizados para encontrar quais as variáveis que mais afetam um determinado processo e a interação entre elas (DE SOUZA; POPPI, 2012). Uma das características mais vantajosas dos métodos quimiométricos é a possibilidade de quantificação ou classificação sem a necessidade de resolução do sinal analítico. Assim, a separação física ou química dos interferentes pode ser substituída pela separação matemática. Essa característica tem proporcionado para o desenvolvimento de vários métodos analíticos simples, rápidos e parcialmente baratos que demanda pouco pré-tratamento das amostras, não precisam de reagentes ou solventes e não geram resíduos, contribuindo com o meio ambiente (TIBOLA et al., 2018).

3.3.1 Análise dos componentes principais (PCA)

A análise de componentes principais (*Principal Component Analysis*, PCA) é um dos métodos mais importantes utilizados na quimiometria e, cada vez mais, novos usuários do meio acadêmico e da indústria estão interessados em sua utilização. Trata-se de um método exploratório que ajuda na elaboração de hipóteses a partir dos dados coletados e permite a redução da dimensionalidade através da representação do conjunto de dados em um novo sistema de eixos, denominados componentes principais (*Principal Component*, PC), permitindo a visualização do comportamento multivariado dos dados em poucas dimensões. Essa análise exerce a capacidade de se separar informações importantes das irrelevantes (DE SOUZA; POPPI, 2012).

Os resultados são analisados frequentemente através do gráfico de escores e pesos, que fornecem informações úteis sobre tendências entre as amostras e variáveis, respectivamente. Esses gráficos podem permitir a visualização de características da estrutura dos dados, como: agrupamento de amostras ou variáveis e caracterização de *outliers*. Além disso, a PCA é a base para diversos métodos de reconhecimento de padrões, classificação e calibração multivariada (DE SOUZA; POPPI, 2012; MARCELO M. DE SENA, RONEI J. POPPI, ROSA T.S. FRIGHETO, 1999).

As áreas que mais se destacam na quimiometria são: reconhecimento de padrões, planejamento de experimentos e calibração multivariada. No reconhecimento de padrões, o objetivo é encontrar grupos de amostras que são similares entre si, a partir de uma grande quantidade de informações sobre vários objetos. Já na calibração multivariada, busca-se estabelecer um modelo que relacione uma série de medidas realizadas em amostras com uma determinada propriedade (DE SOUZA; POPPI, 2012).

A calibração multivariada é uma ferramenta muito útil para a extração de informações química de espectroscopia e tem como princípio básico a utilização simultânea de muitas variáveis, como valores de absorbância ou de comprimentos de onda, para quantificar alguma outra variável de interesse, como, por exemplo, a concentração. (FEUDALE et al., 2002; MARCELO M. DE SENA, RONEI J. POPPI, ROSA T.S. FRIGHETO, 1999).

3.3.2 Pré-processamentos e Ortogonalização de parâmetros externos

O tratamento dos espectros brutos, ou pré-processamento é necessário para a construção dos modelos quimiométricos e seu principal objetivo é eliminar as informações não relevantes de uma matriz, permitindo uma melhor análise exploratória de um conjunto de dados (DE SOUZA; POPPI, 2012).

Dentre os pré-processamentos, podemos destacar a primeira derivada com *Savitzky-Golay* e o *mean center*. Os filtros de alisamento de *Savitzky-Golay* são também conhecidos como alisamento polinomial ou filtros de alisamentos de mínimos quadrados (ORFANIDIS, SOPHOCLES, 2010). São usados normalmente para suavizar um sinal com ruídos e cuja amplitude de frequência é alta. Esses filtros são eficazes na preservação de componentes de sinal de alta frequência e diminuem o efeito prejudicial sobre a relação sinal-ruído que as derivadas convencionais fazem (RINNAN; BERG; ENGELSEN, 2009). O *mean center* centra os dados na média e é realizado através de um cálculo médio de todos os espectros de todas as linhas em um conjunto de dados, subtraindo ponto por ponto de cada espectro desses dados. Graficamente, a centralização da média corresponde a uma mudança em sua origem, permitindo que as diferenças nas intensidades relativas sejam mais facilmente percebidas (DE SOUZA; POPPI, 2012; GEMPERLINE, 2006).

No método de ortogonalização de parâmetros externos (EPO), os espectros são pré-processados através de uma projeção ortogonal, visando remover ou minimizar as alterações induzidas pelas variações do parâmetro externo considerado. A base teórica do método pressupõe que o espaço definido pelos dados espectrais \mathbf{x} é composto pela soma de dois subespaços, entre os quais apenas um contém informações úteis para o modelo e

outro devido às variações de um parâmetro externo. O subespaço devido ao parâmetro externo pode ser estimado de duas maneiras, encontrando o espaço ortogonal às informações de interesse ou encontrando o espaço em que decorre da influência de fatores externos (ROGER; CHAUCHARD; BELLON-MAUREL, 2003).

Em outras palavras, a EPO decompõe um espectro em dois componentes: um útil que tem uma relação direta com a variável de resposta e um que não é útil, proveniente de uma influência externa. Esse componente, devido a influência do parâmetro externo, é estimado através de um conjunto de amostras em que o espectro foi medido em duas condições distintas. Estudos realizados com amostras de solo mostram que modelos desenvolvidos com os espectros corrigidos pela EPO se mostram menos sensíveis à influência do parâmetro externo corrigido e mais preciso quando aplicado a novas amostras (WIJEWARDANE; GE; MORGAN, 2016). MINASNY et al aplicou o algoritmo EPO para remover o efeito da umidade do solo dos espectros NIR. Os resultados mostraram que o algoritmo identificou as áreas dos espectros afetadas pela umidade e puderam remover essas informações (MINASNY et al., 2011).

3.3.3 Regressão por mínimos quadrados parciais (PLSR)

O modelo de regressão em componentes principais (Partial Least Squares Regression, PLSR), é o método mais utilizado na calibração multivariada. A PLSR relaciona duas matrizes de dados, \mathbf{x} e \mathbf{y} , e suas estruturas em um modelo multivariado linear. Para a construção de um modelo, a PLS retira informações do conjunto de dados de uma matriz espectral \mathbf{x} para correlacioná-las com as informações retiradas do conjunto de dados de referência \mathbf{y} . Esse modelo tem a capacidade de analisar dados que apresentam muitas variáveis com ruídos. A PLSR é realizada para cada propriedade de interesse, que utiliza o método de mínimos quadrados entre uma matriz de dados instrumentais \mathbf{x} referente a propriedade de interesse \mathbf{y} . (MORGANO et al., 2008; WOLD; SJÖSTRÖM; ERIKSSON, 2001).

O PLS pode ser uma ferramenta útil quando existe um conhecimento parcial dos dados e os modelos podem ser muito robustos desde que as amostras que serão obtidas futuramente contenham características semelhantes com os dados originais. Um dado importante sobre o PLS é que os erros nas estimativas de concentração e espectros são levados em consideração. Grande parte da história do PLS na química analítica refere-se a aplicações na espectroscopia NIR, em que existem problemas específicos como questões de linhas de base (BRERETON, 2000).

Uma abordagem conhecida é chamada de PLS1 em que ao invés de modelar exclusivamente variáveis \mathbf{x} , são obtidos dois conjuntos de modelo. Sua extensão foi sugerida há 15 anos e é chamada de PLS2. Há pouca diferença conceitual entre os dois, o 1 precisa de um modelo para cada composto e no 2 todos os compostos conhecidos podem ser incluídos em um só modelo. Em alguns casos o PLS2 pode ser mais fácil de se usar computacionalmente pois a análise do modelo é feita de uma só vez, porém, em algumas situações as estimativas são inferiores às do PLS1 (BRERETON, 2000).

4 MATERIAIS E MÉTODOS

4.1 Espécies de madeiras estudadas

Para as análises foram selecionadas amostras do cerne (parte comercial da madeira) do lenho de quatro espécies: *Swietenia macrophylla* King *nativa* (mogno); *Cedrela odorata* L. (cedro); *Micropolis melinoniana* (curupixá) e *Erisma uncinatum* (cedrinho), coletadas e depositadas na Xiloteca Harry Van der Sloten do Laboratório de Produtos Florestais (LPF), Brasília, DF, adquiridas no comércio local.

4.2 Preparação do material

Para o experimento, 30 amostras de cerne de cada espécie foram utilizadas, totalizando 120 amostras. Em seguida, as tábuas foram cortadas nas dimensões 8 x 5 x 1 cm em sentido longitudinal (tangencial ou radial) e lixadas com lixa nº 80 para homogeneizar a superfície da madeira e minimizar a reflexão especular. O excesso de pó de madeira foi retirado com o auxílio de um pincel. A Figura 1 evidencia as amostras preparadas para a aquisição dos espectros. A identificação das espécies das amostras foi confirmada por um dos anatomistas de madeira do LPF.

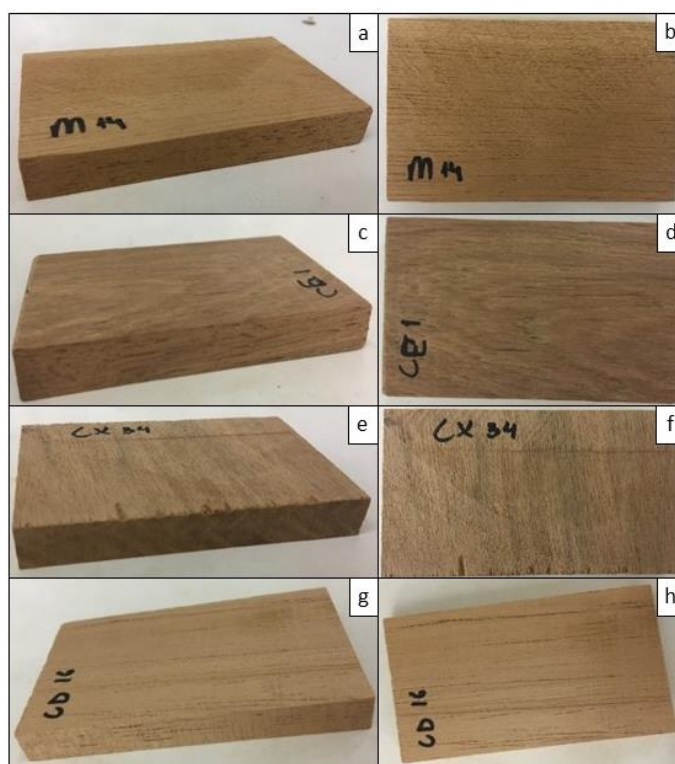


Figura 1 – Amostras de madeira cortadas e lixadas, sendo a) e b) mogno (M), c) e d) cedrinho (CE), e) e f) curupixá (CX), g) e h) cedro (CD). (Fonte: o autor).

4.3 Aquisição dos espectros NIR

Três equipamentos de mesma marca e modelo foram utilizados: espectrômetro portátil microPhazir™ RX Analyzer (Thermo Scientific, Boston, MA, EUA) (Figura 2), que obtém espectros de reflectância difusa na faixa de 1.595 - 2.396nm e resolução de 8,7nm. O equipamento 1 foi adquirido em 2014 e os equipamentos 2 e 3, em 2018.

Para cada uma das espécies estudadas foram medidos espectros com cada um dos equipamentos disponíveis. Os espectros foram obtidos em triplicata em pontos distintos de cada amostra. Os dados foram transferidos para um computador para que se pudesse realizar as análises dos dados e a construção dos modelos de predição no próprio software disponibilizado pelo fabricante do equipamento.



Figura 2 - espectrômetro portátil microPhazir™ RX Analyzer (Thermo Scientific) (Fonte: SCIENTIFIC, 2018)

4.4 Processamento dos espectros e construção de modelos quimiométricos

Os espectros obtidos foram inicialmente submetidos a Análise dos Componentes Principais (PCA, do inglês *Principal Component Analysis*) com e sem pré-processamento.

Os pré-processamentos utilizados foram: Padronização Normal de Sinal (SNV, do inglês *Standard Normal Variate*), primeira derivada com o algoritmo Savitzky-Golay, *Mean Center* e *Normalize Range*. Tanto para as PCAs quanto para os pré-processamentos nesses testes iniciais, foram utilizados os programas Matlab R2017a e PLS Toolbox 7.03.

Para o desenvolvimento dos modelos de predição de cada espécie em cada equipamento está esquematizado na figura 3. Os espectros das amostras foram divididos em dois grupos: treinamento e validação. O grupo de treinamento continha espectros de 20

amostras com três replicatas (total de 60 espectros) de cada espécie; enquanto o grupo de validação contém espectros de 10 amostras com três replicatas (total de 30 espectros) de cada espécie. A seleção dos espectros foi feita de forma aleatória. No total foram utilizados 90 espectros para cada espécie estudada. Já para o desenvolvimento do modelo global, juntou-se espectros dos três equipamentos de todas as espécies estudadas para desenvolver um único modelo de treinamento. Para isso, foram selecionadas, ao acaso, 7 amostras de treinamento de cada espécie e equipamento.

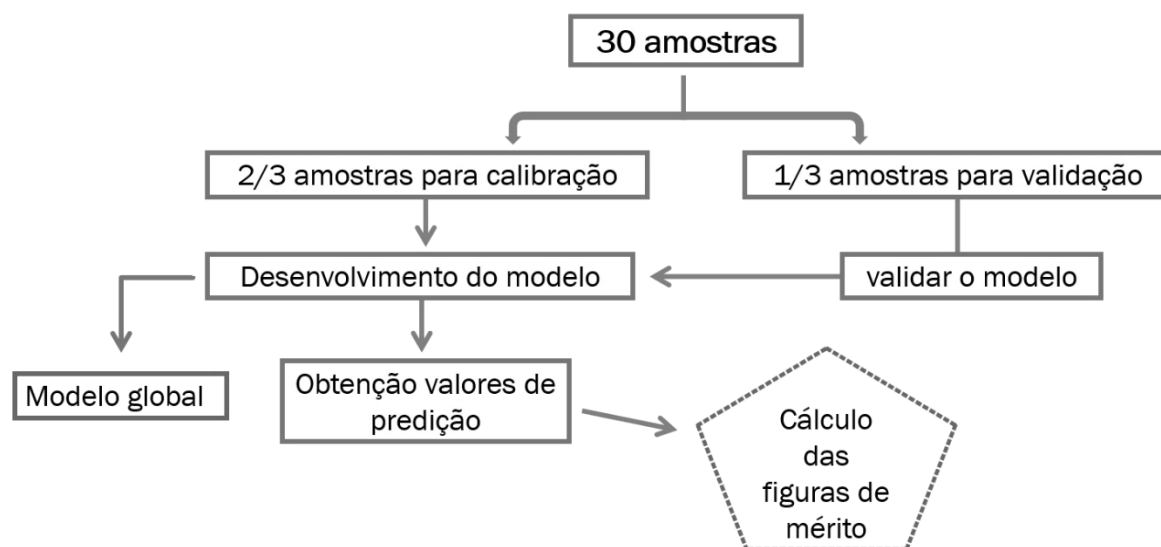


Figura 3 – Procedimento realizado na montagem dos modelos de discriminação. (Fonte: o autor)

A partir dos grupos de calibração foram construídos modelos de predição baseados na Regressão por Mínimos Quadrados Parciais (PLSR, do inglês *Partial Least Squares Regression*) com o auxílio do software do próprio equipamento, Method Generator. Durante esta etapa de treinamento foi feita uma primeira avaliação da performance do modelo a partir do cálculo dos erros médios, expressos em termos da raiz quadrada do erro médio quadrático de calibração (RMSEC, do inglês *root mean square error of calibration*) e do erro médio da validação cruzada (RMSECV, do inglês *root mean square error of cross-validation*) (OLIVEIRA; BRAGA; COSTA, 2015).

A validação cruzada é uma técnica para estimar a performance do modelo na qual o cálculo do PLS é feito retirando espectros ou grupos de espectros. Os valores de classe estimados dos espectros deixados de fora pelo processo de validação cruzada são então previstos pelo modelo criado e comparados com os valores de classe de referência conhecidos desses espectros. O RMSECV é, então, um estimador do desempenho do modelo (SCIENTIFIC, 2018a).

No desenvolvimento dos modelos foi necessário a aplicação do pré-processamento de alisamento e primeira derivada (RINNAN; BERG; ENGELSEN, 2009). Para o alisamento, utilizou-se uma janela de 5 pontos e ajuste de polinômio de segunda ordem, tanto para a PCA quanto para o PLSR. Para realizar a discriminação da espécie de interesse, tanto para treinamento quanto para validação, foi atribuído o valor de referência de classe 1 e para as outras espécies o valor de classe 0.

Os espectros do grupo de validação foram aplicados aos modelos de treinamento construídos. Para a validação, foram testados espectros originados em cada um dos equipamentos para cada modelo de treinamento. A fim de se verificar o percentual de classificação correta com os dados de validação.

As figuras de mérito utilizadas para avaliar a qualidade dos modelos foram: taxas de eficiência do método (TEF), taxa de erros falso positivo (TFP) e erros falso negativo (TFN) descritos abaixo, de acordo com (BOTELHO, B. G, REIS, N, OLIVEIRA, L. S SENA, 2015).

- i. Taxa de falso positivo (TFP): baseia-se na porcentagem de espectros que apresentam erros falso positivos, ou seja, é a porcentagem de espectros de madeira de outras espécies que foram classificados como pertencentes à espécie de interesse. O cálculo é realizado pela razão do número de falsos positivos (FP) pela soma do número de erros falsos positivos (FP) e verdadeiros negativos (VN), multiplicado por 100.

$$TFP = \frac{FP}{FP+VN} 100 \quad (\text{equação 1})$$

- ii. Taxa de falso negativo (TFN): baseia-se na porcentagem de espectros que apresentam erros falsos negativos, ou seja, é a porcentagem de espectros de madeira pertencentes à espécie de interesse que foram classificados como sendo madeira de outras espécies. O cálculo é realizado em relação ao número de falsos negativos (FN) pela soma do número de erros falsos negativos (FN) e verdadeiros positivos (VP) multiplicado por 100.

$$TFN = \frac{FN}{FN+VP} 100 \quad (\text{equação 2})$$

- iii. Taxa de eficiência (TEF): Representa a eficácia do modelo desenvolvido na discriminação das espécies desejadas. Pode ser calculada pela equação:

$$TEF = 100 - (TFN + TFP) \quad (\text{equação 3})$$

As figuras de mérito foram calculadas com auxílio de uma planilha eletrônica, empregando o software Excel.

5 RESULTADOS E DISCUSSÃO

5.1 Análise exploratória dos espectros obtidos com os três equipamentos empregando PCA

A Figura 4 apresenta a distribuição gráfica bidimensional dos espectros obtidos para o mogno em cada equipamento estudado a partir das novas variáveis obtidas pelas PCA, antes e após o pré-processamento dos espectros. Em cada gráfico, pode ser observado os escores de cada componente principal, sendo que a PC1 descreve a máxima variância dos dados originais e a PC2, apresenta a direção de máxima variância dos dados no subespaço ortogonal à PC1. Os espectros dos equipamentos 1, 2 e 3 estão representados pelos losangos vermelhos, quadrados verdes e triângulos azuis, respectivamente. A mesma análise exploratória realizada por PCA para as outras espécies estudadas, cedrinho, cedro e curupixá, estão apresentados nas figuras 11, 12 e 13, respectivamente, em anexo.

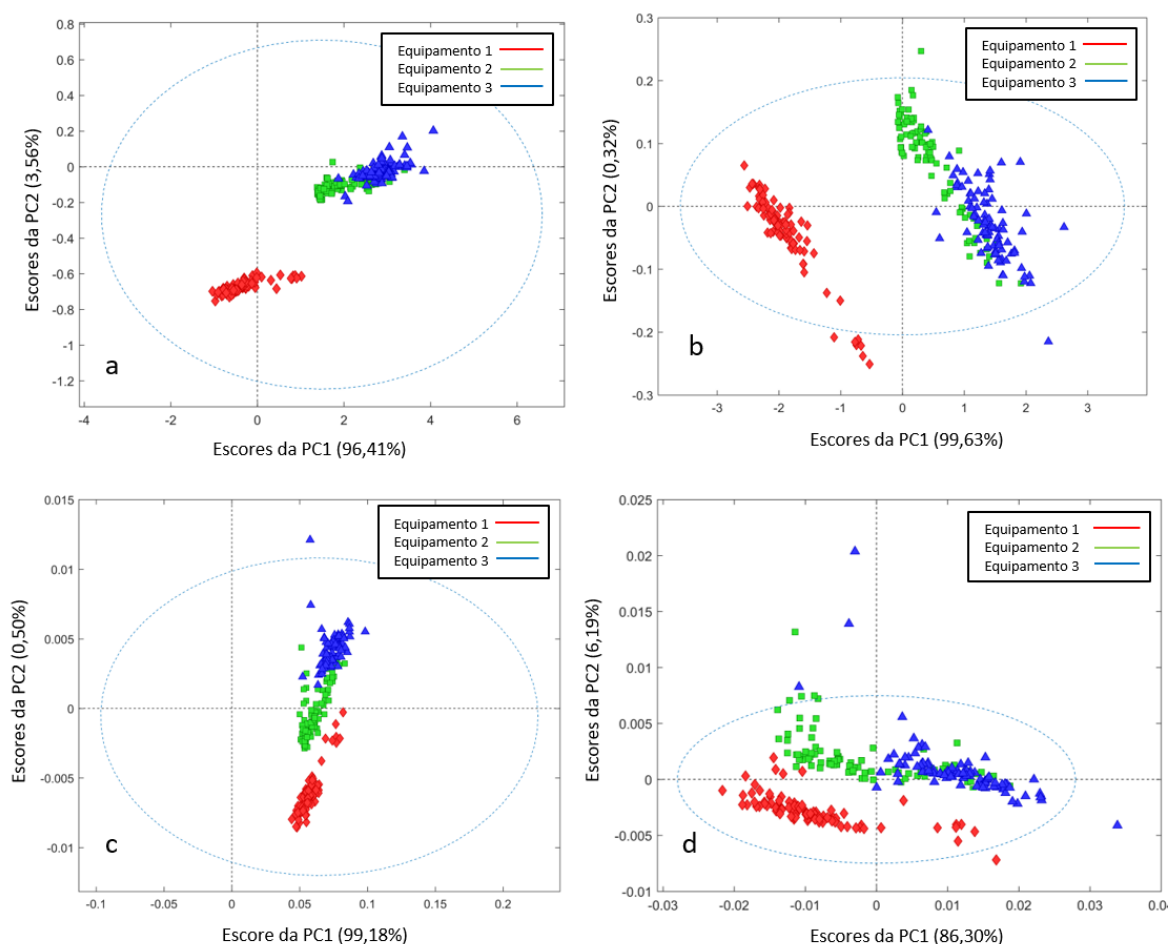


Figura 4 - Gráficos bidimensionais dos espectros de mogno a partir da PCA dos três equipamentos estudados, antes e após o pré-processamento dos espectros: (a) sem pré-processamento; (b) pré-processamento mean center; (c) pré-processamento 1ª derivada com Savitzky-Golay e (d) junção dos pré-processamentos mean center e 1ª derivada (Fonte: o autor).

Pode-se observar que os espectros brutos, ou seja, sem aplicação de pré-processamentos, originados do equipamento 1 agrupam-se de forma distinta dos espectros originados pelos equipamentos 2 e 3, que, por outro lado formaram nuvens que parcialmente sobrepostas (Fig. 4a). Ao se aplicar o primeiro pré-processamento referente ao *mean center*, foi possível uma melhor visualização da sobreposição parcial dos espectros dos equipamentos 2 e 3, enquanto os espectros do equipamento 1 permanecem isolados (Fig. 4b). Por outro lado, após aplicar a primeira derivada como pré-processamento, observou-se uma aproximação das nuvens de pontos de cada equipamento (Fig. 4c). Essa aproximação dos espectros ficou ainda mais evidente ao se aplicar a primeira derivada associada ao *mean center* (Fig. 4d). Este fato também ocorreu para as demais espécies estudadas e suas respectivas figuras estão apresentadas no Anexo. Em todos os casos, PC1 e PC2 juntas explicaram mais de 90% dos dados.

A forma de agrupamento dos espectros brutos evidenciadas na Fig. 4a pode ser explicada pela diferença nas intensidades da radiação de infravermelho emitida por cada um dos equipamentos. Tal hipótese é suportada pelos resultados observados na Fig. 5, no qual os grupos espectrais obtidos apresentaram valores de linha de base que variam conforme o equipamento. Nesta figura, o equipamento 1 gerou espectros com linhas de base bem mais baixas do que os espectros dos equipamentos 2 e 3. Como resultado, assim como foi observado nos gráficos de escores, os espectros brutos do equipamento 1 ficaram bastante afastados dos espectros dos outros dois espectrômetros.

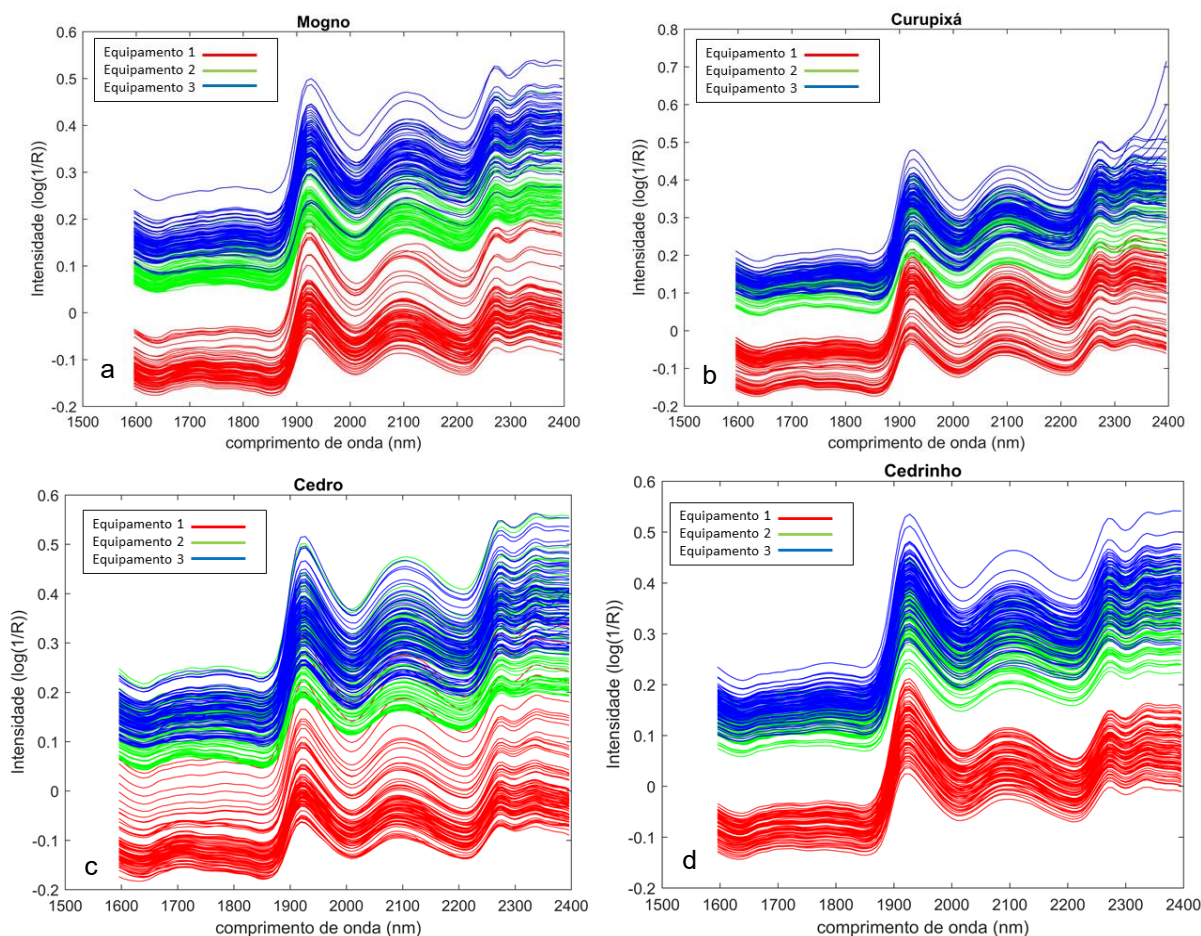


Figura 5 – Espectros brutos obtidos dos três equipamentos para as espécies: (a) mogno; (b) curupixá; (c) cedro e (d) cedrinho (Fonte: o autor).

É possível que equipamentos semelhantes possam fornecer respostas instrumentais diferentes, como a alteração no deslocamento de linha de base no comprimento de onda, gerando mudanças nas respostas espectrais medidas (HONORATO et al., 2007). De acordo com os autores, existem basicamente três causas que reduzem a eficiência de um modelo em casos de variações na etapa de calibração: mudanças na constituição química e/ou física das amostras, mudanças no ambiente do equipamento ou mudanças na função de resposta do equipamento, que farão com que o espectro de uma amostra apresente o mesmo perfil, porém, com intensidades diferentes. Essa mudança relacionada ao equipamento pode ser causada pelo envelhecimento da fonte e/ou do detector.

Assim, ainda que possam existir pequenas diferenças entre intensidades luminosas entre equipamentos de mesmo modelo, as diferenças de intensidades de reflectância observadas estão provavelmente relacionados com os desgastes da fonte do equipamento 1, que apresenta maior tempo de uso que os equipamentos 2 e 3. O equipamento 1 foi

adquirido pelo LPF em 2014, com bem mais horas de uso que os equipamentos 2 e 3, adquiridos em 2018.

Na tentativa de remover variabilidades indesejáveis nos espectros causados por efeitos de espalhamento ou variações entre equipamentos, e ao mesmo tempo para preservar a informação de interesse, podem ser aplicados os pré-processamentos (HONORATO et al., 2007).

A figura 6 evidencia as alterações promovidas pela aplicação da primeira derivada com o filtro de alisamento de Savitzky-Golay nos espectros dos três equipamentos para as quatro espécies estudadas. As diferenças observadas para os espectros brutos tornaram-se menos perceptíveis, sugerindo que este pré-tratamento reduziu as variações espectrais causadas pelas diferenças dos equipamentos. A redução dessas diferenças foi, também, provavelmente a causa para a sobreposição das nuvens espectrais observadas para a PCA da Fig. 4d. De forma geral, a análise por PCA pode revelar que os espectros de cada equipamento apresentam diferenças sutis, que fazem com que estes se agrupem por equipamento para todas as amostras analisadas.

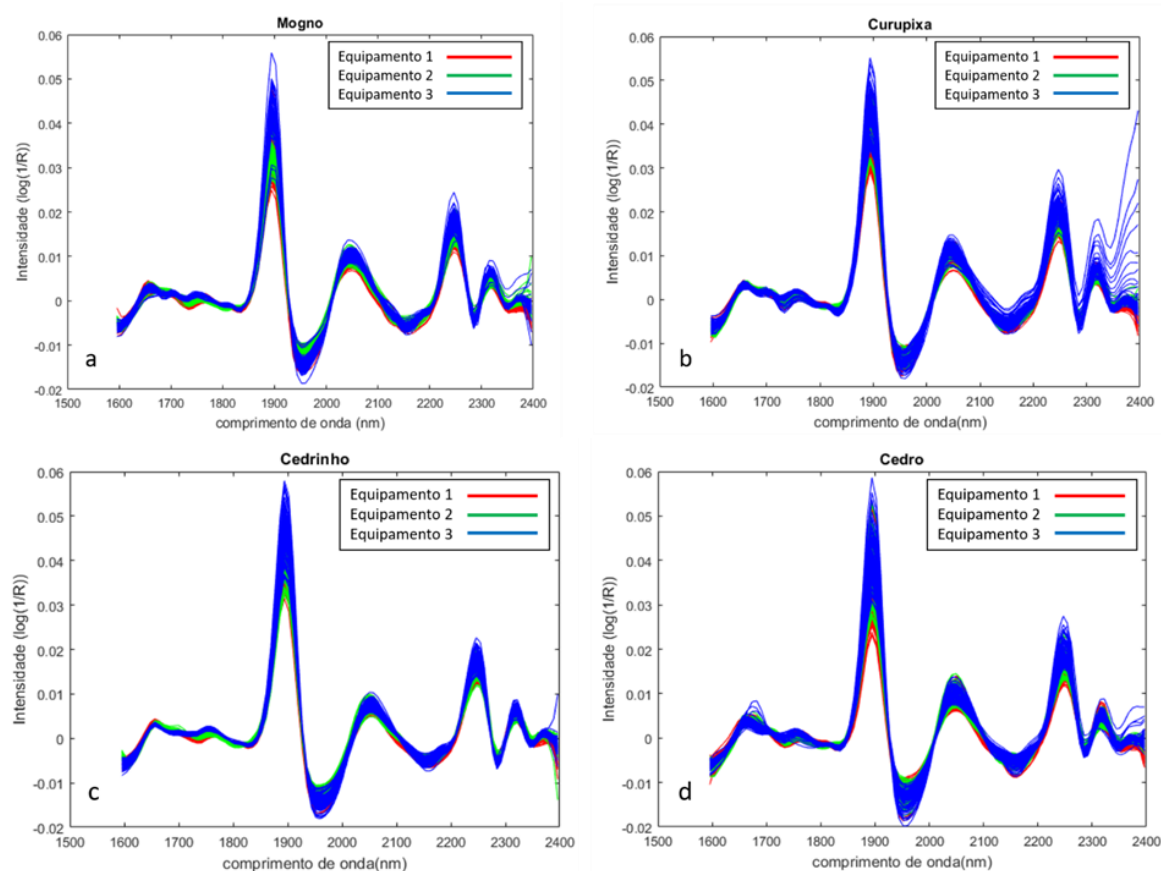


Figura 6 – Espectros brutos obtidos dos três equipamentos após a aplicação da 1ª derivada com alisamento Savitzky Golay das espécies: (a) mogno; (b) curupixá; (c) cedrinho e (d) cedro. (Fonte: o autor)

5.2 Modelo de predição

Os resultados das figuras de mérito dos modelos de predição baseados em PLSR de cada espécie em cada equipamento estão apresentados na tabela 1, que contém os resultados dos modelos calibrados com espectros originados nos equipamentos 1, 2 e 3 (Modelo 1, Modelo 2 e Modelo 3, respectivamente) validados com espectros obtidos pelos dos equipamentos 1, 2 e 3 (Validação 1, Validação 2 e Validação 3, respectivamente).

De acordo com a Tabela 1, as maiores taxas de eficiência foram obtidas ao se validar espectros de um equipamento no modelo de discriminação obtido com os espectros do próprio equipamento de origem, sendo que houve uma tendência próxima ou igual a 100% nesses casos. Essas taxas estão marcadas na tabela em azul. A exceção foi observada para o cedrinho cuja TEF, ao se validar as amostras do espectrômetro 2 com o modelo de calibração do espectrômetro 1, foi de 97,5%. Ao se validar um modelo de calibração com espectros originados de outro equipamento, as TEFs tenderam a uma queda drástica, entre 0 e 67,5% e estão marcadas em vermelho. Este resultado já era esperado de acordo com as diferenças espectrais observadas por PCA entre os equipamentos.

Tabela 1 - Valores das figuras de mérito para os modelos de discriminação para os equipamentos 1, 2 e 3 e o modelo global.

	Mogno								
	Validação 1			Validação 2			Validação 3		
	TEF	TFP	TFN	TEF	TFP	TFN	TEF	TFP	TFN
Modelo 1	100%	0%	0%	57,5%	42,5%	0%	0%	0%	100%
Modelo 2	0%	0%	100%	95%	5%	0%	10%	0%	90%
Modelo 3	10%	0%	90%	10%	0%	90%	100%	0%	0%
Modelo global (mix)	82,5%	17,5%	0%	75%	25%	0%	70%	20%	10%
	Cedrinho								
	Validação 1			Validação 2			Validação 3		
	TEF	TFP	TFN	TEF	TFP	TFN	TEF	TFP	TFN
Modelo 1	100%	0%	0%	97,5%	2,5%	0%	90%	0%	10%
Modelo 2	100%	0%	0%	100%	0%	0%	90%	0%	10%
Modelo 3	20%	0%	80%	70%	0%	30%	100%	0%	0%
Modelo global (mix)	90%	0%	10%	90%	0%	10%	100%	0%	0%

	Cedro								
	Validação 1			Validação 2			Validação 3		
	TEF	TFP	TFN	TEF	TFP	TFN	TEF	TFP	TFN
Modelo 1	100%	0%	0%	20%	0%	80%	55%	5%	40%
Modelo 2	20%	80%	0%	40%	0%	60%	50%	0%	50%
Modelo 3	0%	0%	100%	85%	15%	0%	100%	0%	0%
Modelo global (mix)	100%	0%	0%	100%	0%	0%	100%	0%	0%
	Curupixá								
	Validação 1			Validação 2			Validação 3		
	TEF	TFP	TFN	TEF	TFP	TFN	TEF	TFP	TFN
Modelo 1	100%	0%	0%	100%	0%	0%	92,5%	7,5%	0%
Modelo 2	100%	0%	0%	100%	0%	0%	67,5%	32,5%	0%
Modelo 3	0%	0%	100%	0%	0%	100%	100%	0%	0%
Modelo global (mix)	90%	0%	10%	90%	0%	10%	87,5%	2,5%	10%

TEF: Taxa de eficiência; TFP: Taxa de falso positivo; TFN: Taxa de falso negativo; os números 1, 2 e 3 referem-se aos equipamentos utilizados.

Nota-se também que o resultado das TEFs dos modelos variou conforme a espécie. Tomando como exemplo as espécies mogno e cedrinho, observa-se que as taxas de cedrinho são mais altas que as taxas de mogno. Este fato pode ser explicado devido a diferença espectral entre as espécies, na qual o cedrinho costuma ser discriminado mais facilmente do que as demais espécies do que o mogno, independentemente da condição de modelo x validação submetida.

A figura 6 apresenta, para diferentes situações da tabela 1, a distribuição dos valores estimados de classe obtidos tanto para o conjunto de treinamento quanto para o de validação, além dos limites de discriminação (LD) estabelecidos para os valores da espécie de mogno.

Para uma discriminação da espécie de interesse eficiente, os valores estimados tendem a 1 e para as demais espécies esses valores tendem a 0, sem que haja sobreposição entre os agrupamentos ou centros de distribuição (DA SILVA et al., 2014).

O gráfico **a** (Fig. 7a) representa a validação do equipamento 1 no modelo de discriminação do mesmo equipamento 1. Já o gráfico **b** (Fig. 7b), mostra a validação do equipamento 2 no modelo de discriminação do equipamento 3, e no gráfico **c** (Fig. 7c) apresenta a validação do equipamento 2 no modelo global de discriminação. A classe 1 corresponde aos espectros de mogno e classe 0 os espectros das demais espécies estudadas.

Pelos gráficos, nota-se que o valor do LD foi aproximadamente menor que 0,5 e maior que 0,4. No gráfico **a** observa-se que acima do LD tem-se a espécie de interesse mogno discriminada e abaixo as demais espécies estudadas, evidenciando a TEF de 100%. Em **b**, observa-se que as amostras de validação de mogno estão abaixo LD, resultando em uma TEF de 10% e na TFN de 90%. Por fim, em **c** as amostras de validação das demais espécies estudadas estão acima do LD, resultando numa TEF de 75% e na TFP de 25%.

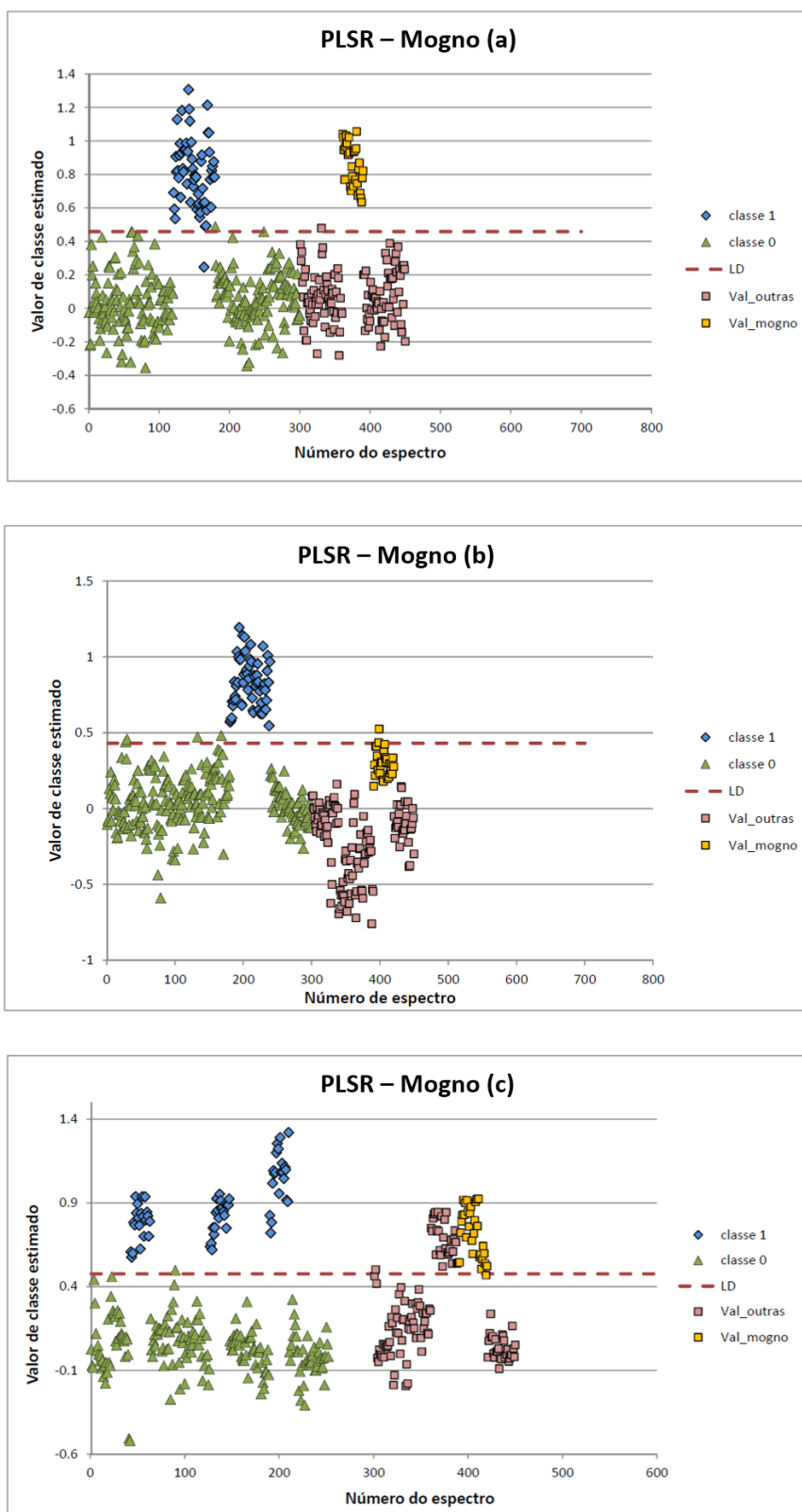


Figura 7 – Distribuição dos valores estimados do conjunto de treinamento e validação para os modelos PLSR para o mogno. a) calibração eq. 1 com validação eq. 1; b) calibração eq. 3 com validação eq. 2; c) calibração modelo global com validação eq. 2; Classe 1 = mogno; Classe 0 = demais espécies estudadas; LD = limite de discriminação; Val_outras = espectros de validação das demais espécies e Val_mogno = espectros de validação do mogno. (Fonte: o autor)

Como uma forma de aumentar as TEFs entre as combinações de treinamento/validação de diferentes equipamentos de origem, foi desenvolvido um modelo de calibração a partir da junção dos espectros das quatro espécies estudadas dos três equipamentos (modelo global ou mix), também representados na tabela de resultados 1 e marcado na cor rosa. Pode-se observar que as TEFs das validações dos modelos globais foram superiores a aquelas cujas validações foram feitas com modelos de discriminação com espectros de outro equipamento de origem.

De acordo com HONORATO et al. (2007), quando há aplicação de espectros de um modelo desenvolvido em um determinado equipamento (primário) em um outro equipamento (secundário), é provável que haja diferenças entre as respostas instrumentais, mesmo que o equipamento seja um só, pois sua resposta pode variar ao longo do tempo. Consequentemente, dois equipamentos podem fornecer diferentes respostas instrumentais. Como forma de reduzir tais diferenças, indica-se a padronização, na qual a função de resposta do equipamento secundário é modificada para se tornar parecida com a função de resposta do equipamento primário, onde o modelo foi originalmente desenvolvido.

No presente trabalho, ao invés de se recorrer ao método de padronização, optou-se pelo desenvolvimento do modelo global, considerado mais parcimonioso e de menor complexidade, como estratégia para melhorar as taxas de eficiência e melhorar o desempenho de discriminação das espécies entre equipamentos de mesmo modelo.

5.3 Avaliação preliminar da correção da diferença entre equipamentos por EPO

Uma das medidas tomadas para tentar equalizar ou aproximar as respostas instrumentais dos equipamentos, foi realizar uma correção preliminar à análise exploratória através da EPO, ortogonalização de parâmetros externos (do inglês *External Parameter Orthogonalization*), em que o parâmetro externo que foi corrigido está relacionado a diferença entre os espectros dos três equipamentos. Como o equipamento 2 e 3 não possuem diferenças espectrais tão evidentes, a análise preliminar da correção com EPO foi realizada entre os espectros obtidos com os equipamentos 1 e 2, utilizando os espectros da madeira mogno, visto que é a espécie de maior interesse deste trabalho.

A figura 8a mostra os espectros brutos após a aplicação da 1ª derivada com alisamento Savizky Golay e, na figura 8b, os espectros corrigidos com a aplicação da EPO.

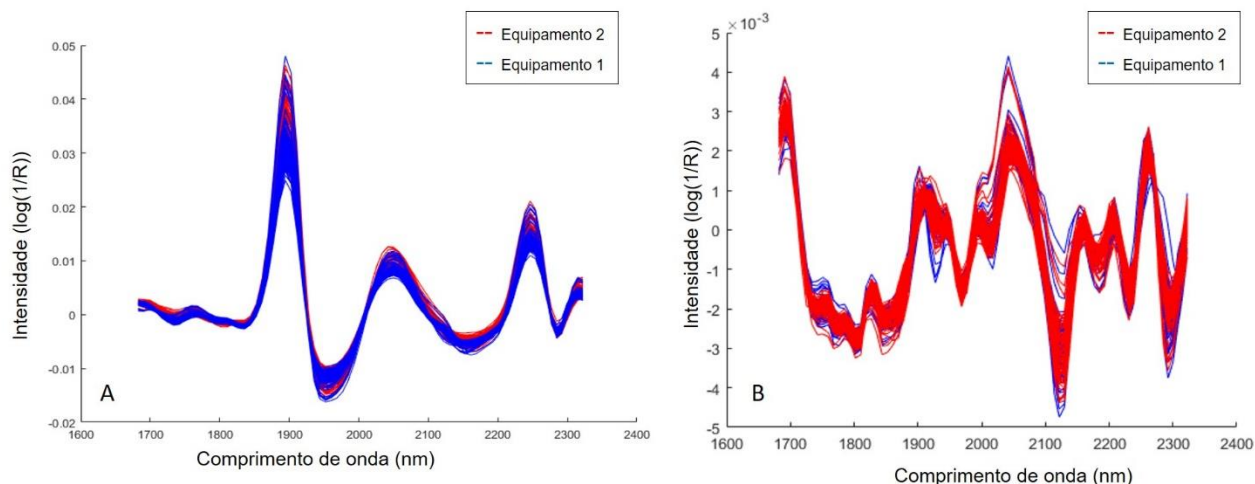


Figura 8 – Espectros brutos de mogno dos equipamentos 1 e 2 em que: (A) aplicação da 1ª derivada com alisamento Savizky Golay e (B) correção com EPO. (Fonte: o autor).

A partir disso, a PCA, descrita no tópico 5.1, foi aplicada aos espectros com e sem a correção para analisar se houve uma aproximação dos agrupamentos dos espectros dos equipamentos 1 e 2, a fim de aproximar as diferenças observadas anteriormente. A figura 9 apresenta o gráfico de escores da PCA em que foi aplicado o *mean center* como pré-processamento sem a aplicação da EPO. Tem-se, então, em **A** o gráfico de duas PCs e em **B** o gráfico com três PCs. O resultado após aplicação da EPO está evidenciado na figura 10. Assim, da mesma forma, em **A** tem-se o gráfico de duas PCs e em **B** o gráfico com três PCs.

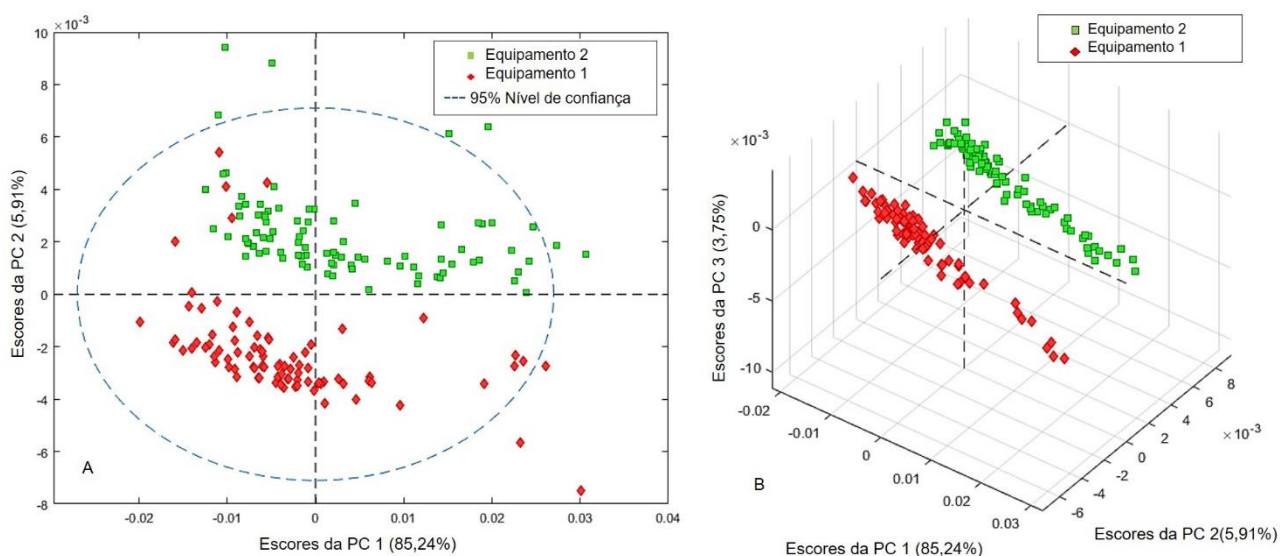


Figura 9 - Gráficos bidimensionais dos espectros de mogno a partir da PCA dos equipamentos 1 e 2 com o pré-processamento *mean center* e sem a aplicação da EPO: (A) gráfico com duas PCs e (B) gráfico com três PCs (Fonte: o autor).

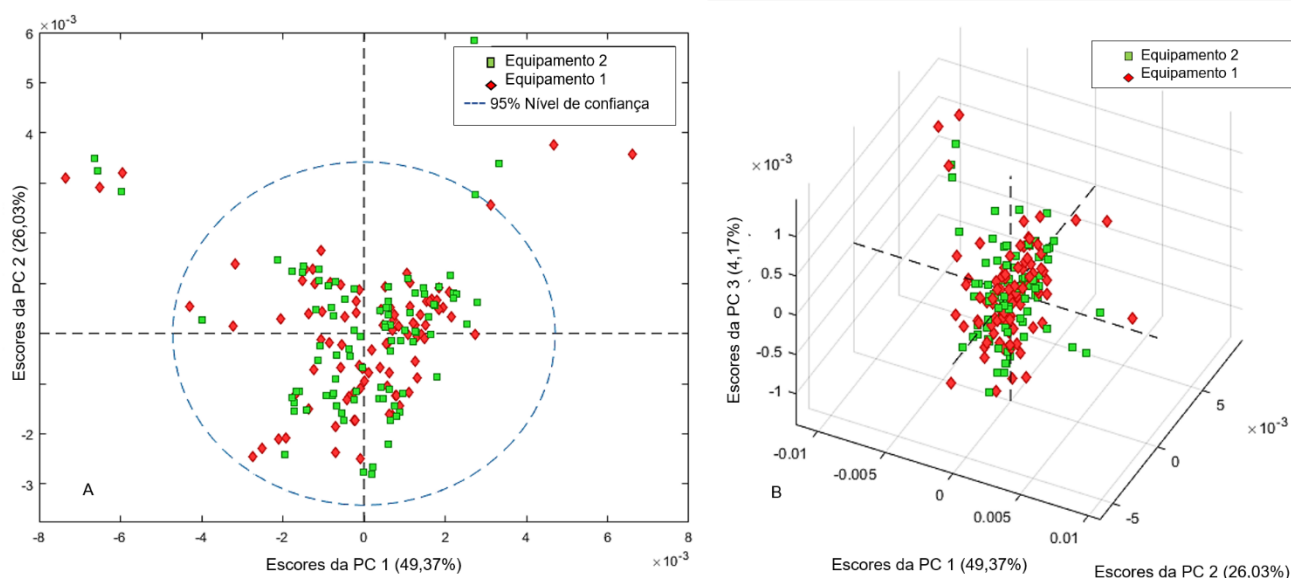


Figura 10 - Gráficos bidimensionais dos espectros de mogno a partir da PCA dos equipamentos 1 e 2 com o pré-processamento *mean center* e com a aplicação da EPO: (A) gráfico com duas PCs e (B) gráfico com três PCs (Fonte: o autor).

A partir da figura 10 nota-se o resultado da correção entre as diferenças espectrais dos equipamentos 1 e 2, em que houve a sobreposição das nuvens, ou seja, a aproximação dos espectros. Esse fato evidencia que a utilização da EPO pode corrigir essa diferença entre os espectros, não havendo mais uma diferença entre os equipamentos estudados e, portanto, será, provavelmente capaz de aumentar as taxas de eficiência dos modelos de discriminação quando testados.

6 CONCLUSÕES E PERSPECTIVAS

A partir dos resultados apresentados, pode-se concluir que existem diferenças nos equipamentos, apesar de serem da mesma marca e modelo ou com mesmo tempo de uso. Entretanto, tais diferenças podem ser minimizadas a partir da aplicação de pré-processamento, realizando a construção de modelos de calibração que combinem espectros obtidos em diferentes equipamentos e aplicando a EPO como correção preliminar.

A EPO foi a estratégia que mais aproximou as respostas instrumentais dos equipamentos analisados. Assim, ajustes nos espectros brutos para a construção dos modelos podem ser realizados não apenas para se utilizar mais de um equipamento ao mesmo tempo, mas também para corrigir eventuais desgastes do equipamento que possam impactar em sua eficiência.

Como perspectivas, a partir dos resultados obtidos neste trabalho, sugere-se carregar os modelos obtidos utilizando a correção com a EPO em cada um dos equipamentos a fim de se testar amostras externas e comparar os resultados de identificação em cada um deles. Sugere-se ainda aumentar o número de amostras utilizadas para que se possa ter uma melhor representação das espécies e assim, tornar os modelos mais robustos.

7 REFERÊNCIAS BIBLIOGRÁFICAS

BERGO, M. C. J. **Transferência de Calibração na discriminação de mogno e espécies semelhantes utilizando NIRS e PLS-DA**. Brasília-DF: Universidade de Brasília, 2014.

BERGO, M. C. J. et al. NIRS identification of *Swietenia macrophylla* is robust across specimens from 27 countries. **IAWA Journal**, v. 37, n. 3, p. 420–430, 2016.

BOTELHO, B. G, REIS, N, OLIVEIRA, L. S SENA, M. M. Development and analytical validation of a screening method for simultaneous detection of five adulterants in raw milk using mid-infrared spectroscopy and PLS-DA. **Food Chemistry**, v. 181, n. 31, 2015.

BRASIL. Decreto Lei nº 4.722, de 5 de junho de 2003. **Diário Oficial da União, Poder Executivo**, 2003.

BRASIL. Decreto Lei nº 6.472, de 5 de junho de 2008. **Diário Oficial da União, Poder Executivo**, 2008.

BRASIL. Portaria Ministério do Meio Ambiente. **Portaria Ministério do Meio Ambiente**, v. número 443, 2014.

BRERETON, R. G. Introduction to multivariate calibration in analytical chemistry. **The Analyst**, v. 125, n. 11, p. 2125–2154, 2000.

CORADIN, V. T. R.; CAMARGOS, J. A. A. MARQUES, L. F.; SILVA JÚNIOR, E. R. DA S. Madeiras similares ao mogno (*Swietenia macrophylla* King.): chave ilustrada para identificação anatômica em campo. p. 28, 2009.

DA SILVA, V. A. G. et al. Non-destructive identification of different types and brands of blue pen inks in cursive handwriting by visible spectroscopy and PLS-DA for forensic analysis. **Microchemical Journal**, v. 116, p. 235–243, 2014.

DE MUNIZ, G. I. B. et al. Fundamentos e estado da arte da espectroscopia no infravermelho próximo no setor de base florestal. **Ciência Florestal**, v. 22, n. 4, p. 865–875, 2012.

DE SOUZA, A. M.; POPPI, R. J. Experimento didático de quimiometria para análise exploratória de óleos vegetais comestíveis por espectroscopia no infravermelho médio e análise de componentes principais: UM tutorial, parte I. **Química Nova**, v. 35, n. 1, p. 223–229, 2012.

ESBENSEN, K.; GELADI, P. The start and early history of chemometrics: Selected interviews. Part 1. **Journal of Chemometrics**, v. 4, n. 5, p. 337–354, 1990.

FEUDALE, R. N. et al. Transfer of multivariate calibration models: A review. **Chemometrics and Intelligent Laboratory Systems**, v. 64, n. 2, p. 181–192, 2002.

GEMPERLINE, P. **Practical Guide to Chemometrics**. 2^a edition ed. Broken: [s.n.].

HONORATO, F. A. et al. Transferência de Calibração em Métodos Multivariados. **Química Nova**, v. 30, n. 5, p. 1301–1312, 2007.

MACIEIRA, A. P. et al. Caracterização morfoanatômica de mogno (*Swietenia macrophylla* King - MELIACEAE) para fins de identificação botânica. **Anais do XII Seminário Anual de Iniciação Científica da UFRA**, p. 0–3, 2013.

MARCELO M. DE SENA, RONEI J. POPPI, ROSA T.S. FRIGHETO, P. J. V. Avaliação do uso de métodos quimiométricos em análise de solos. **Química Nova**, v. 23, n. 4, p. 531–537, 1999.

MINASNY, B. et al. Removing the effect of soil moisture from NIR diffuse reflectance spectra for the prediction of soil organic carbon. **Geoderma**, v. 167–168, p. 118–124, 2011.

MORGANO, M. A. et al. Determinação de umidade em café cru usando espectroscopia NIR e regressão multivariada. **Ciencia e Tecnologia de Alimentos**, v. 28, n. 1, p. 12–17, 2008.

OLIVEIRA, A. A. et al. Identificação de Madeiras utilizando a Espectrometria no Infravermelho Próximo e Redes Neurais Artificiais. **TEMA (São Carlos)**, v. 16, n. 2, p. 81, 2015.

OLIVEIRA, E. M.; BRAGA, J. W. B.; COSTA, A. F. DA. Discriminação de madeiras similares por fluorescência molecular e mínimos quadrados parciais. **Química Nova**, v. 38, n. 9, p. 1176–1180, 2015.

ORFANIDIS, SOPHOCLES, J. **Introduction to Signal Processing**. [s.l: s.n.].

PASQUINI, C. Near infrared spectroscopy: Fundamentals, practical aspects and analytical applications. **Journal of the Brazilian Chemical Society**, v. 14, n. 2, p. 198–219, 2003.

PASTORE, T. C. M. et al. Near infrared spectroscopy (NIRS) as a potential tool for monitoring trade of similar woods: Discrimination of true mahogany, cedar, andiroba, and curupixá. **Holzforschung**, v. 65, n. 1, p. 73–80, 2011.

RINNAN, Å.; BERG, F. VAN DEN; ENGELSEN, S. B. Review of the most common pre-processing techniques for near-infrared spectra. **TrAC - Trends in Analytical Chemistry**, v. 28, n. 10, p. 1201–1222, 2009.

ROGER, J. M.; CHAUCHARD, F.; BELLON-MAUREL, V. EPO-PLS external parameter orthogonalisation of PLS application to temperature-independent measurement of sugar content of intact fruits. **Chemometrics and Intelligent Laboratory Systems**, v. 66, n. 2, p. 191–204, 2003.

SCIENTIFIC, T. **Method Generator Software - Instruction Manual 4.3**. [s.l.] Thermo Scientific, 2018a.

SCIENTIFIC, T. F. **microPHAZIR Analyzer**, 2018b.

SILVA, D. C. et al. Determination of the country of origin of true mahogany (*Swietenia macrophylla* King) wood in five Latin American countries using handheld NIR devices and multivariate data analysis. **Holzforschung**, v. 72, n. 7, p. 521–530, 2018.

SKOOG, D. A.; HOLLER, F. JAMES NIEMAN, T. A. **Princípios de Análise Instrumental**. 5ª edição ed. [s.l.] Bookman, [s.d.].

SO, B. C. et al. Near infrared spectroscopy in the forest products industry. **Forest Products Journal**, v. 54, n. 3, p. 6–16, 2004.

SOARES, L. F. et al. Avaliação de espectrômetro NIR portátil e PLS-DA para a discriminação de seis espécies similares de madeiras amazônicas. **Química Nova**, v. XY, n. 00, p. 1–9, 2017.

SOUZA, M. H. DE; CAMARGOS, J. A. A. Madeiras Tropicais Brasileiras. **Serviço Florestal Brasileiro**, v. 2, n. LPF, p. 156, 2014.

TIBOLA, C. S. et al. (EDS.). **Espectroscopia no Infravermelho Próximo para Avaliar Indicadores de Qualidade Tecnológica e Contaminantes em Grãos**. Brasília - DF: Embrapa, 2018.

TSUCHIKAWA, S. A review of recent near infrared research for wood and paper. **Applied Spectroscopy Reviews**, v. 42, n. 1, p. 43–71, 2007.

WIJEWARDANE, N. K.; GE, Y.; MORGAN, C. L. S. Moisture insensitive prediction of soil properties from VNIR reflectance spectra based on external parameter orthogonalization. **Geoderma**, v. 267, p. 92–101, 2016.

WOLD, S.; SJÖSTRÖM, M.; ERIKSSON, L. PLS-regression: A basic tool of chemometrics. **Chemometrics and Intelligent Laboratory Systems**, v. 58, n. 2, p. 109–130, 2001.

8 ANEXO

Relação de figuras do resultado da PCA das espécies estudadas antes e após o pré-processamento dos espectros: (a) sem pré-processamento; (b) pré-processamento mean center; (c) pré-processamento 1ª derivada com Savizky-Golay e (d) junção dos pré-processamentos mean center e 1ª derivada.

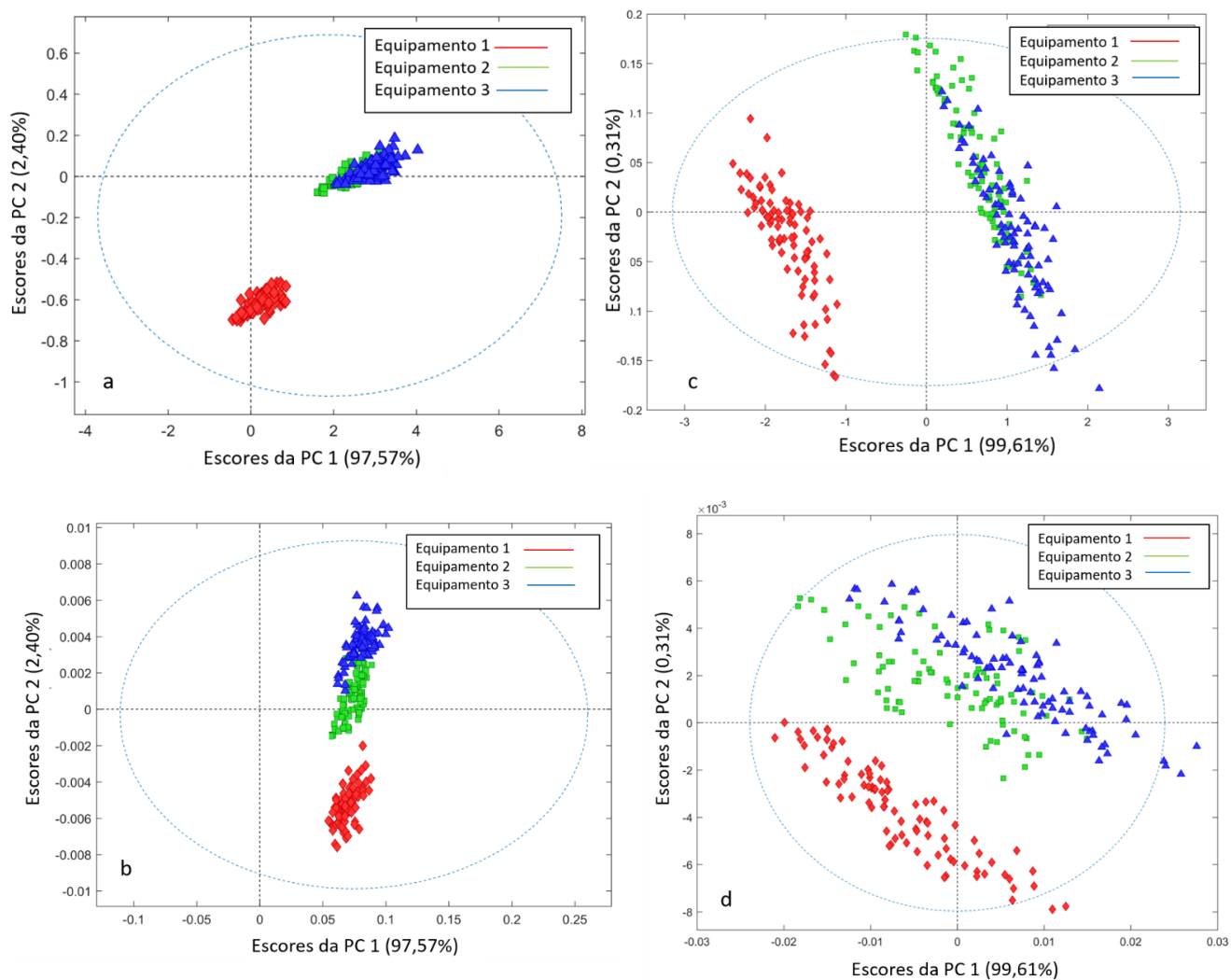


Figura 11 - Gráficos bidimensionais dos espectros de cedrinho (Fonte: o autor).

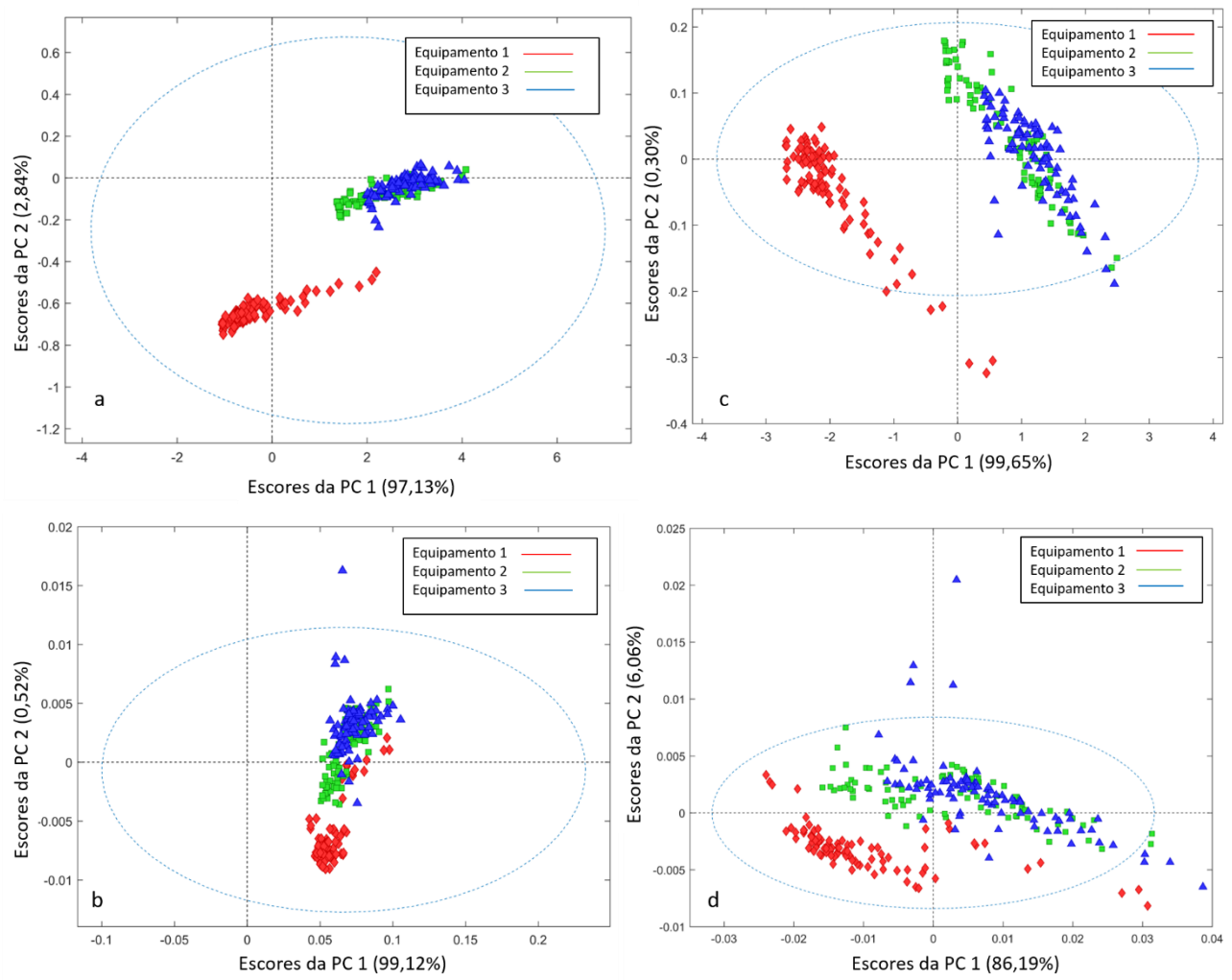


Figura 12 - Gráficos bidimensionais dos espectros de cedro (Fonte: o autor)

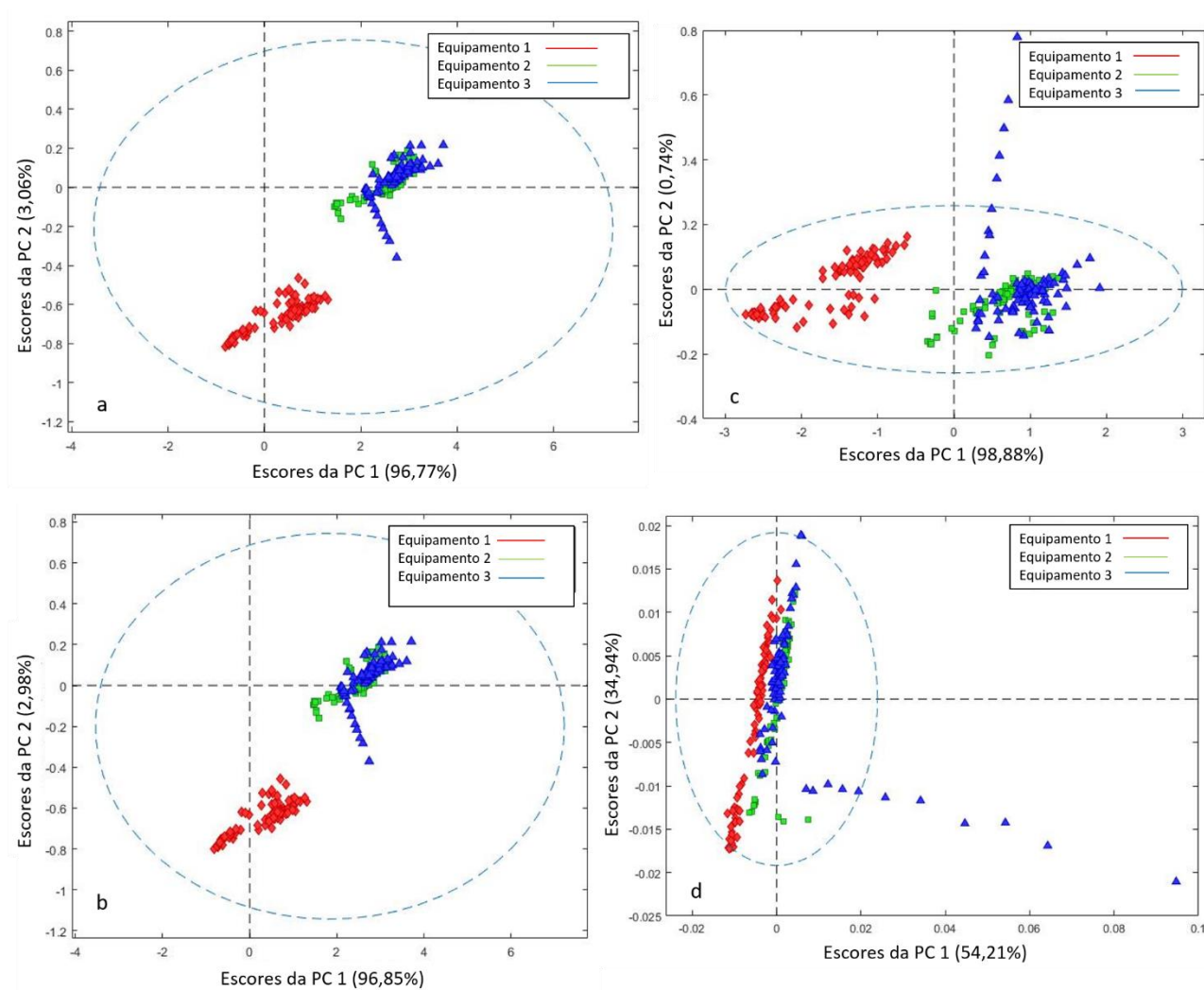


Figura 13 - Gráficos bidimensionais dos espectros de curupixá (Fonte: o autor).